Genet Program Evolvable Mach (2012) 13:537–538
DOI 10.1007/s10710-012-9168-3

BOOK REVIEW

Alma Lilia Garcia Almanza, and Edward Tsang: Evolutionary applications for financial prediction: classification methods to gather patterns using genetic programming

VDM Verlag Dr. Müller, 2011, ISBN: 978-3639307672

Blake LeBaron

Published online: 29 June 2012

It is nice to see a book applying evolutionary algorithms to the problem of financial forecasting. This is one of the very few books on the market that spells out in detail an evolutionary system as applied to financial time series. It is a detailed book, and concentrates on many of the computational details of building an evolutionary forecasting system with applications to some financial data sets along with some standard machine learning benchmarks.

The book begins with several introductory chapters. First, there is a chapter on machine learning from a classification perspective. It emphasizes the problem of imbalanced classes, which is a running theme throughout the book. (By imbalanced classes we mean that the numbers of positive and negative cases are dramatically different.) Dealing with these kinds of data issues is important, and is understudied in the field of finance. Also, some metrics are presented for measuring the usefulness of a learning tool. Particularly important is something known as a ROC (receiver operating characteristic) curve, that can help users judge a forecast based on their own attitude toward false positives in a classification system. Chapter 3 provides a succinct, but adequate summary of methods in evolutionary computation. Readers slightly familiar with these methods will probably find it useful, but it is not a complete tutorial on GP or evolutionary computation in general. That would require a lot of space.

There is a brief summary of financial ideas in chapter 4. Here, I think the survey is a little weak. It provides only a light overview of financial markets, and most of the chapter is spent describing various technical trading indicators used for forecasting. There is no commentary on what sorts of objectives traders might have, or on any of the institutional details of the markets that are going to be analyzed. Anyone without a good finance background should not expect this book to teach

B. LeBaron (✉)
International Business School, Brandeis University, Waltham, MA, USA
e-mail: blebaron@brandeis.edu

them this. The spirit of staying relatively light on ideas from finance continues throughout much of the book.

The authors then set up three different strategies for classification systems that are explored in the later chapters. The repository method (RM) concentrates on building an ensemble of classifiers which can be chosen by the user according to their classification preferences over false positives, and true case identification. The evolutionary decision rule (EDR) system works from a population of rules and continues to evolve it through time. Generally, this process is well described, but I found the exact statement about fitness to be a little unclear. Given that much of the work here stresses how the usefulness of a classification rule depends on user's preferences, it is important how these preferences would map into the evolutionary fitness criterion for rule selection. Finally, a system known as the Scenario Method is applied. The Scenario Method takes a population of rules and carefully prunes each. This appeared to be generally successful at improving classification.

All of the classification methods are applied to financial and nonfinancial data. Comparisons are made using the ROC test, and some of its related tests. Garcia Almanza and Tsang appear to be using distinct testing and training sets, but this is not made clear in some of the examples. Also, some of the meta parameters are presented without information on how they were chosen. In these sections some more direct comparisons across methods would have been useful. For some of the comparisons the differences across the methods appear to be relatively small, and given the noisiness in financial data, I'm not sure they are significant.

My biggest concern is that the authors should have tried some more standard financial forecasting tests. This would include setting these classifiers up as a trading rule, and seeing how well they do. Good predictive systems can sometimes not do so well when implemented as a dynamic trading strategy. I came away with no finance measure for the significance of these results. I think the concept of classifying unbalanced samples is important, but this may generate unusual results as a trading system. Rules that concentrate on avoiding large losses may end giving up too much on the gain side. Again, only a trading simulation would be able to tell you exactly what to expect.

It is refreshing to see a detailed book on evolutionary computation in finance. Finance professionals who are interested in seeing some of this technology in action may find this book useful. They should take special note of the emphasis on unbalanced sampling, which is under researched in the area of finance. The overall results of the *Evolutionary Applications for Financial Practitioners* leave me a little cautious, since none of the tests can really tell how well a rule would stand up in an actual trading situation. Computer scientists interested in finance may find this a useful example of an application to look at. However, this book should not be relied on to provide the necessary background in finance to develop a trading system.